MOLECULAR MACHINERY: A Tour of the Protein Data Bank

Living cells are filled with complex molecular machinery, a million times smaller than familiar machines like computers or automobiles. Cells use these tiny molecular machines to perform all of the jobs needed for life. Some are molecular scissors that cut food into cell-sized pieces. Some build new molecules when cells grow or when damaged tissues are repaired. Some are molecular bones and muscles that support cells and help them move and crawl. Some fight off attackers, defending against infection.

Researchers around the world are studying these molecules and determining their precise atomic structures. These structures are available on the internet through the Protein Data Bank (http://www.pdb.org), the central storehouse of biomolecular structures. A few of the thousands of structures held in the Protein Data Bank are shown here. In these pictures, the molecules are all drawn at a magnification of 3,000,000 times, and each atom is shown as a small sphere. Many of these structures are composed of several subunits, which are indicated by different colors. An enormous range of sizes is shown here: the water molecule at the left has only three atoms and the rhinovirus shown below has hundreds of thousands.

Water
Glucose
Adenosine Triphosphate (ATP)

By David S. Goodsell, The Scripps Research Institute, La Jolla, California, USA

Graphic design by Gail W. Bamber, San Diego Supercomputer Center

1 OUTSIDE THE CELL

Some molecular machines perform their jobs outside of cells. Many are compact, so that they can diffuse quickly to their site of action. This is true of the four hormones shown at the top: insulin and glucagon, which together regulate blood sugar levels, interferon, which carries signals in the immune system, and human growth hormone. The seven digestive enzymes (in yellow) are also small and very stable, so that they can survive the hostile environment in the digestive tract. Each of these enzymes has a small groove (oriented towards the top in each) that binds to a different target molecule and digests it. At the bottom is rhinovirus, the virus that causes the common cold, and an antibody, our major defense against viruses. Antibodies bind to viruses and prevent them from binding to cell surfaces, thus blocking infection.

1gcn Glucagon
2hiu Insulin
3hhr Human Growth Hormone
1rfb Interferon
2ptc Trypsin
2dnj Deoxyribonuclease
1lz1 Lysozyme
1poe Phospholipase
1smd Amylase
5rsa Ribonuclease
5pep Pepsin
1igt Antibody
4rhv Rhinovirus

Cholesterol
Phospholipid
1c17•1e79 ATP Synthase
1bgy Cytochrome b-c1 Complex
3cyt Cytochrome c
1oco Cytochrome c Oxidase
1f88 Rhodopsin
1pth Cyclooxygenase
1prc Photosynthetic Reaction Center
1kzu Light-Harvesting Complex
1jb0 Photosystem I

2 MEMBRANES

Cells are surrounded by a membrane made of lipids, like the phospholipid and cholesterol molecules shown at the top. Membranes keep the cellular machinery inside and unwanted material out. Many proteins are embedded in this membrane, performing a variety of essential tasks. ATP synthase is a rotary generator that produces ATP (adenosine triphosphate), the small molecule used for powering cells. The two large complexes below it charge a battery that powers ATP synthase, and the tiny protein cytochrome c shuttles electrons between them. Rhodopsin is found in membranes in the retina. The small retinol molecule inside of it changes shape when illuminated, causing the surrounding protein to send a signal to the brain. Cyclooxygenase builds one of the molecules used to signal pain—this cyclooxygenase molecule here, however, is blocked by two molecules of aspirin, shown inside in white. At the bottom are three molecules involved in photosynthesis, which capture energy from light and use it to power the synthesis of sugar in plant cells.

1fqy Aquaporin
1bl8 Potassium Channel
1msi Mechanosensitive Channel
1grm Gramicidin
2por Porin
4hhb Hemoglobin
1mbd Myoglobin
1hrs Ferritin
1e7i Serum Albumin

3 TRANSPORT AND STORAGE

Of course, a perfectly sealed membrane would be of little use to cells, because nutrients could not get in and wastes could not get out. The box shows a membrane looking face-on. Five proteins that form channels through the membrane are shown. To the right of the box are several soluble proteins involved in transport and storage of molecules. Hemoglobin and myoglobin carry oxygen. Ferritin forms a hollow shell that stores iron ions. Serum albumin carries many different molecules in the blood.

1cza Hexokinase
1iat Phosphoglucose Isomerase
1pfk Phosphofructokinase
1ald Aldolase
7tim Triosephosphate Isomerase
3gpd Glyceraldehyde-3-phosphate Dehydrogenase
3pgk Phosphoglycerate Kinase
3pgm Phosphoglycerate Mutase
5enl Enolase
1a49 Pyruvate Kinase
1dhf Dihydrofolate Reductase
2ohx Alcohol Dehydrogenase
1wsy Tryptophan Synthase
2tsc Thymidylate Synthase
2gls Glutamine Synthetase
4at1 Asparate Carbamoyltransferase
1rcx Ribulose Bisphosphate Carboxylase/Oxygenase
1n2c Nitrogenase

4 CHEMICAL FACTORIES

Cells build a bewildering variety of enzymes—proteins that perform chemical reactions. At the top are the ten enzymes that perform glycolysis, the breakdown of sugar to form ATP. Below that are several enzymes that perform different housekeeping reactions. Dihydrofolate reductase activates a key cofactor molecule and alcohol dehydrogenase breaks down alcohol. Ribulose bisphosphate carboxylase/oxygenase is the most common enzyme on the Earth, and performs a key step in the capture of carbon dioxide by plants to form sugar. The three synthases and the transferase make different building blocks for creating new molecules. Nitrogenase performs an essential role in the ecosystem by converting nitrogen gas into a form that living cells can use.

1cj9 Topoisomerase
1i6h RNA Polymerase
1ais TATA-binding Protein/ Transcription Factor IIB
1cgp Catabolite Gene Activator Protein
1tau DNA Polymerase
1bkh•1efa lac Repressor
1aoi Nucleosome
1aoi Nucleosome
1aon Chaperonin GroEL/ES

5 DNA

Genetic information is stored in the DNA double helix, seen running from top to bottom here. Many proteins are used to copy, read, and store this information. RNA polymerase copies the information into a strand of RNA that will be used to direct the construction of new proteins. It is assisted by topoisomerase, which releases tension when the helix is wound and unwound, and guided to appropriate starting points by the two protein complexes below it. DNA polymerase replicates DNA strands—here, the polymerase is filling a gap in the double helix. Some proteins, like the lac repressor, grab DNA and bend it sharply, or even wrap it all the way around themselves, like the two nucleosomes at the bottom.

1gax Valyl-tRNA Synthetase
1qf6 Threonyl-tRNA Synthetase
1ffy Isoleucyl-tRNA Synthetase
1eiy Phenylalanyl-tRNA Synthetase
1asy Aspartyl-tRNA Synthetase
1euq Glutaminyl-tRNA Synthetase
1ffk•1jj2 Ribosome
4tna Transfer RNA
1ttt Elongation Factor Tu
1efu Elongation Factor Tu and Ts
1dar Elongation Factor G
2cpl Proline cis/trans Isomerase
1fxk Prefoldin

6 BUILDING NEW PROTEINS

New proteins are built by ribosomes—complex molecular factories that read the genetic code and use it to direct construction. Many molecular machines are needed to assist in the process. Twenty different aminoacyl-tRNA synthases (six are shown here) load the building blocks onto tRNA, ready to be added to a growing protein chain. Several protein factors, shown below the ribosome, guide each tRNA into the proper spot. The three chaperone proteins shown at the bottom help each new protein fold into its proper shape.

1b7t Myosin
1tub Microtubule
1atn Actin
1bkv Collagen

7 BEAMS AND GIRDERS

Cells are braced and supported by a complex infrastructure. This cytoskeleton is formed of sturdy filaments like actin and microtubules, composed of many subunits stacked like bricks. Myosin is a molecular motor that climbs along actin filaments, allowing the cell to move. Collagen, broken into two pieces here, is actually found outside of cells, where it forms connective tissue between cells.

February 2002

PROTEIN DATA BANK

http://www.pdb.org/ • info@rcsb.org

RESEARCH COLLABORATORY FOR STRUCTURAL BIOINFORMATICS

RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY
SAN DIEGO SUPERCOMPUTER CENTER
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

• **PDB ID: 1a49** E.M. Larsen, M.M. Banning, I. Raymont, G.H. Reed (1998) Structure of the bis(Mg2+)-ADP-oxalate complex of the rabbit muscle pyruvate kinase at 2.1 A resolution: ADP binding over a barrel. *Biochemistry* **37**, p. 4247. • **PDB ID: 1ais** P.D. Kim, L. Ghosh, B.S. Dalbeiker, P.E. Taylor (1997) The 2.1-A crystal structure of an unbound preinitiation complex: TBP/double-looping promoter/transcription factor IIB complex/TBP/dsDNA. *Proc. Natl. Acad. Sci. U.S.A.* **94**, p. 4041. • **PDB ID: 1atn** S.J. Gamblin, G.J. Davies, J.M. Grimes, R.M. Jackson, J.A. Littlechild, N.C. Watson (1991) Activity and specificity of human skeletal muscle actin, bovine cardiac actin, and fish muscle actin in the myosin subfragment 1-activated Mg2+-ATPase reaction. *J. Biol. Chem.* **266**, p. 19644. • **PDB ID: 1atp** J.P. Abrahams, A.G.W. Leslie, R. Lutter, J.E. Walker (1994) Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria determined in three crystal forms. *Nature* **370**, p. 621